Computer model of a bovine type I collagen microfibril

Gregory King, Eleanor M.Brown and James M.Chen[1]

Eastern Regional Research Center, Agricultural Research Service, USDA, Philadelphia, PA 19118, USA

[1]Present address: Du Pont Agricultural Products, Stine-Haskell Research Center, PO Box 30, Building 300, Newark, DE 19714, USA

Collagens are a family of structural proteins of the extracellular matrix. The fibril-forming collagens are the major structural proteins of skin, cartilage, bone, blood vessel walls and internal organs. In addition to biological function, the collagens provide natural structural frameworks that are utilized in the medical, food and leather industries. Many schemes for the organization of type I collagen into triple helices, microfibrils and fibrils have been proposed during the past 30 years. Here, the development of a molecular model of a bovine type I collagen 'Smith' microfibril is described. In cross-section, this model exhibits a symmetrical, pentagonal grouping of five triple helices. The model comprises 15 polypeptide chains having 315 residues each. This model is large enough to allow a comparison of its gross structural features with images of stained collagen obtained by electron microscopy, yet small enough to be manipulated on a minicomputer or workstation. The model is useful for (among others) studies of structure–function relationships in collagen, exploring folding pathways, predicting the efficacy of potential crosslinking agents or chemical modifications, and designing synthetic collagen-like materials or modifications for specific applications.

Keywords: banding pattern/collagen/D space/microfibril/molecular modeling

Introduction

Collagens are a family of structural proteins of the extracellular matrix. At least 14 types of collagen with varying amino acid sequences have been described, each with one or more triple-helical domains (Van der Rest and Garrone, 1991). The fibril-forming collagens, types I, II, III, V and XI, are the major structural proteins of skin, cartilage, bone, blood vessel walls and internal organs. Many studies of collagen structure and function have been directed towards learning the causes of and possible treatments for connective tissue disorders. In addition to biological function, the collagens have commercial importance as natural structural frameworks utilized in the medical, food and leather industries (Bailey, 1992); in the form of gelatin, it can be processed for use in the pharmaceutical, food, glue and photographic industries (Rose, 1992).

Several unique characteristics of the fiber-forming collagens have influenced the choice of technique for exploring details of the molecular structures of these ubiquitous and multipurpose proteins. Because these collagens form fibrillar structures large enough to be visualized using relatively primitive instruments, a picture of the basic organization of collagen in triple-helical, microfibrillar and fibrillar structures began to emerge in the 1960s (Ramachandran, 1967; Smith, 1968). Individual peptide chains are ~1050 amino acid residues long and, except for telopeptides at the N- and C-termini, consist entirely of a repeating Gly–X–Y pattern where at least 25% of X and Y residues are Pro or Hyp (hydroxyproline) respectively. These features of the sequence made the development of template models feasible. Early work on the preparation and use in structural studies of synthetic collagen-like peptides (reviewed by Carver and Blout, 1967) provided baseline data for spectroscopic evaluations of conformations in collagen. The computational model of $(Gly–Pro–Pro)_n$ (Miller and Scheraga, 1976) was the start of a series of related studies which evaluated the effects of specific side chains on conformation. More extensive reviews of these and other studies leading to proposed pathways for the assembly of microfibrils were published at the start of our research (Chen *et al.*, 1991a), including a microfibril model of type II collagen (Chen and Sheldon, 1994).

Type I collagen is the most abundant and its structure has been widely studied. X-ray scattering studies of non-crystalline collagen (Ramachandran, 1968) and a crystallized model peptide (Okuyama *et al.*, 1972) $(Pro–Pro–Gly)_{10}$ synthesized by Sakakibara *et al.* (1968) have verified the hypothesis of Ramachandran and Kartha (1954) that three collagen molecules assemble into triple-helical entities (coiled-coils), such that 27–29 amino acid residues comprise a complete rotation about the right-handed triple-helical axis. The atomic coordinates of the triple-helical peptide $3(Pro–Pro–Gly)_{10}$ were determined by Okuyama *et al.* (1976). The structure of the collagen triple helix is expected to be similar to that of this model peptide, but the details of the packing of triple helices into fibrils are not yet known definitively at the atomic level.

Electron micrographs of stained collagen fibrils (Mould *et al.*, 1990) provide some clues as to the gross packing structure of collagen. Such micrographs display a pattern of alternating light and dark bands perpendicular to the axis of the collagen fibrils that repeat every 670 Å. This 670 Å span has been defined as a D interval (or D spacing). Light bands correspond to regions of more dense lateral packing, and dark bands correspond to 'gap' regions, domains of low density molecular packing first noted by Hodge and Petruska (1963).

Various models for collagen have been proposed based on the observed staining pattern and the length of a single collagen triple-helical molecule 4.4 D intervals in length (Fraser *et al.*, 1974; Veis and Yuan, 1975; Piez and Trus, 1977, 1978; Hofmann *et al.*, 1978; Okuyama *et al.*, 1978; Traub, 1978; Fraser *et al.*, 1979, 1983). It is generally agreed that groups of four to six triple helices are packed together to form microfibrils, which in turn aggregate to form fibrils. The model proposed by Smith (1968) is able to explain much of the electron microscopy data. In this model, a microfibril is defined as a bundle of five triple-helical molecules, in which adjacent triple helices are staggered longitudinally by a 1.0 D interval.

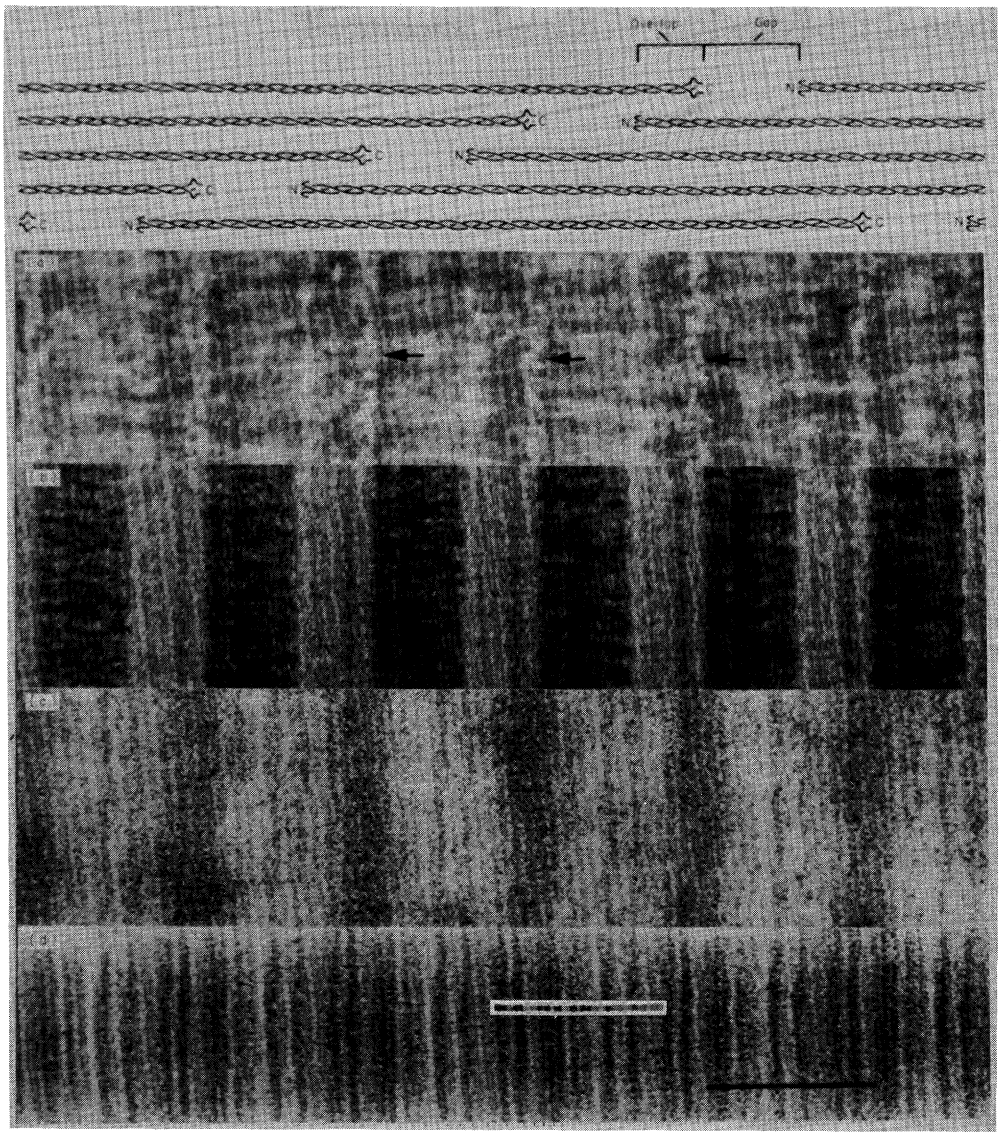

Fig. 1. Photographs of stained chick type I collagen obtained by electron microscopy. The top, unlabeled panel is a cartoon illustrating the staggering of triple helices within collagen fibrils. (**a**) Negatively stained procollagen assembly. (**b**) Negatively stained collagen fibrils. (**c**) Positively stained procollagen assembly. (**d**) Positively stained collagen fibrils. (Reprinted from Mould *et al.*, 1990; Copyright 1990 Academic Press Ltd.) In (b), the negatively stained collagen fibrils display a coarse pattern of alternating light and dark bands, which is thought to be caused by a preferential deposition of stain in regions of the fibril where collagen is present at a lower density (gap regions). More detail is provided in (d), where positive staining is used. The smaller, positively charged stain particles bind to negatively charged regions of the protein. The dark horizontal bar in (d) indicates one D period. For comparison, the banding pattern obtained from our computer model has been superimposed on (d) (a larger-scale version of the computer model's banding pattern is provided in Figure 5).

Longitudinally neighboring triple helices are separated by a 'gap' region ~0.6 D intervals in length (see top panel of Figure 1). Cross-sections of the model exhibit a regular pentagonal geometry.

The choice of the Smith microfibril model as a template for our computer model was influenced by the fact that a complete type I collagen molecule 4.4 D intervals in length (5.0 D intervals when a gap region is included) may be represented with minimal redundancy in a five-molecule microfibril segment of length 1.0 D intervals, caused by the relative longitudinal staggering of the triple-helical molecules by 0.0, 1.0, 2.0, 3.0 and 4.0 D intervals. Some X-ray diffraction studies (e.g. Fraser *et al.*, 1983) indicate that pentagonal groupings of triple-helical molecules are not present in collagen fibrils. The packing of triple-helical molecules prescribed by the Smith microfibril may therefore not be strictly correct when viewed in cross-section. This is not a serious concern because in subsequent studies the model will only be used to investigate crosslinking interactions involving adjacent pairs of triple-helical collagen molecules. The center–center distance of 10.0 Å between adjacent triple-helical molecules in the model compares well with the value of 10.5 Å indicated by X-ray diffraction data for dry collagen fibrils (Rich and Crick, 1961).

In preliminary studies we constructed 36 residue-long template models of the collagen triple helix and microfibril using the template sequences $(Gly–Pro–Pro)_{12}$ and $(Gly–Pro–Hyp)_{12}$ (Chen *et al.*, 1991a), as well as a 36 residue-long segment of bovine type I collagen (Chen *et al.*, 1991b,c). Here, we have used these building blocks in the construction of a much larger computer model of a bovine type I collagen microfibril following the blueprint provided by Smith (1968). This model consists of 15 polypeptide chains, each of length 315 amino

acid residues (1.0 D interval = 234 amino acid residues), and contains the amino acid sequence of bovine type I collagen. The model is thus large enough to be useful in comparing gross structural features with the images of stained collagen obtained by electron microscopy, in predicting the effectiveness of proposed tanning agents for crosslinking collagen molecules to stabilize leather, and in evaluating the probable efficacy of other industrial processes using collagen. At the same time, this model is small enough to be manipulated on a minicomputer or workstation.

Materials and methods

The 3-D computer model of the bovine type I collagen microfibril was constructed using SYBYL molecular modeling software (version 6.0; Tripos Associates, St Louis, MO) on an SGI 4D/35 workstation (Silicon Graphics, Mountain View, CA) and on the Cray supercomputer at the North Carolina State Super Computing Center, NC. Previous studies performed at this laboratory (Chen *et al.*, 1991a,b,c) describe the assembly of computer models of relatively short segments of collagen-like molecules. These molecules are the building blocks that were used to construct the larger collagen model described here. To provide a basis for our model, the construction and use of these building blocks will be reviewed briefly.

In the first stage (Chen *et al.*, 1991a), a single polypeptide chain $(Gly–Pro–Pro)_{12}$ was constructed using the parameters for each amino acid provided by the dictionary component of SYBYL. The conformation of this peptide was initialized by using the backbone dihedral angles ϕ, ψ and ω, which Miller and Scheraga (1976) found in a computational study to give the minimum energy for triple helices of the form $3(Gly–Pro–Pro)_n$. For Gly, the dihedral angles are $\phi = -74°$ and $\psi = 170°$. For the first Pro, the angles are $\phi = -75°$ and $\psi = 168°$. For the second Pro, the angles are $\phi = -75°$ and $\psi = 153°$. The angle ω for all residues is 180°. Next, the triple helix $3(Gly–Pro–Pro)_{12}$ was constructed by visually docking three $(Gly–Pro–Pro)_{12}$ peptide models together. In this triple-helical arrangement, the three peptides are staggered with respect to each other by one residue (an arrangement with efficient packing and strong hydrogen bonds). This structure also possesses one hydrogen bond per amino acid triplet, in agreement with the hydrogen bonding arrangement first proposed for collagen triple helices by Rich and Crick (1961). Then the triple helix was refined via energy minimization, in which the united-atoms AMBER force field (Weiner and Kollman, 1981; Weiner *et al.*, 1984) was used. Next, the model was modified so that every third residue was hydroxyproline, to produce the sequence $3(Gly–Pro–Hyp)_{12}$. This sequence more closely resembles actual collagen sequences, where proline residues in the Y position of Gly–X–Y tend to be hydroxylated; thus fewer modifications were required when the actual type I sequence was substituted into the model at a later stage. Retaining a periodic sequence such as $(Gly–Pro–Hyp)_n$ also made it possible to perform splicing operations to lengthen the molecule (described below) that would not have been feasible if the actual type I sequence was used from the beginning. After these hydroxylations, the model was again energy minimized. Many of the hydroxyl groups of the Hyp residues formed hydrogen bonds with peptide backbone oxygen atoms, but virtually no secondary structural changes were induced. A description of the construction of the model to this point was published by Chen *et al.* (1991a).

In the second stage, Chen *et al.* (1991b,c) constructed and refined by energy minimization five 36 residue-long triple-helical units of type I collagen by substituting amino acid residues from the sequence published by Fietzek and Kuhn (1976) into the X and Y positions of the $(Gly–Pro–Hyp)_{12}$ model. A 36 residue, 15 chain Smith (1968) microfibril model was then constructed by packing five $(Gly–Pro–Hyp)_{12}$ triple helices, each with a right-handed twist, into a left-handed superhelical arrangement. The same portion of the type I sequence used above was then substituted into the microfibril model. These models were refined by energy minimization at each step in their construction. Even these short segments of triple helix and microfibril were useful for visualizing the non-random distribution of charged and hydrophobic side chains predicted by Piez and Trus (1978). These models also proved useful in some preliminary evaluations of potential crosslinking agents for use in leather making (Brown *et al.*, 1992; Scholnick *et al.*, 1992).

Stepwise Transformation of Template Amino Acid Sequence into Actual Collagen Amino Acid Sequence

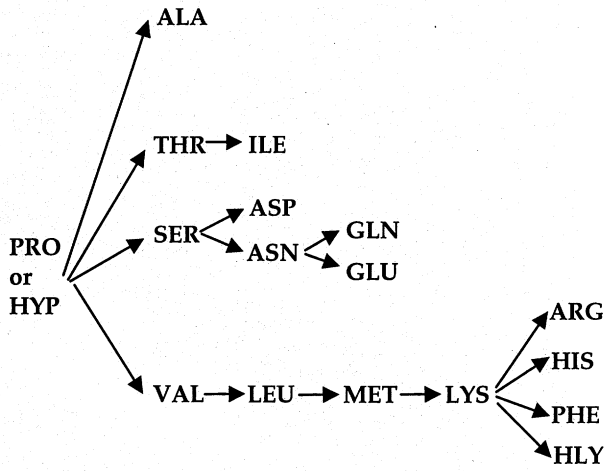

Fig. 2. A flowchart showing the 'alchemical' procedure that was followed when transforming the template sequence $(Gly–Pro–Hyp)_n$ into the actual bovine type I collagen sequence. The final amino acid residues were arrived at through a series of zero to five successive 'mutations' into residues whose side chains increased gradually in size.

To extend the type I collagen microfibril model to include a full D space (234 amino acid residues in length), the following procedure was employed. A splicing operation was used to lengthen the triple-helical $3(Gly–Pro–Hyp)_{12}$ model. In this procedure, two segments $3(Gly–Pro–Hyp)_k$ and $3(Gly–Pro–Hyp)_k$ are aligned with one another so that they overlap slightly (overlapping the two segments makes it easier to visualize whether the two pieces are in register with one another). The overlapping atoms are deleted, and peptide bonds are formed to create one longer entity, $3(Gly–Pro–Hyp)_m$, with m slightly less than $2k$. The splicing procedure was repeated until the model structure $3(Gly–Pro–Hyp)_{69}$ was achieved.

Five copies of the model triple-helical molecule $3(Gly–Pro–Hyp)_{69}$ were bundled together in an arrangement whose cross-section was a regular pentagon to form the microfibril $5[3(Gly–Pro–Hyp)_{69}]$, and the model was energy minimized. The increasing size of the model began to make it difficult to work with on a workstation platform, so the model was transferred

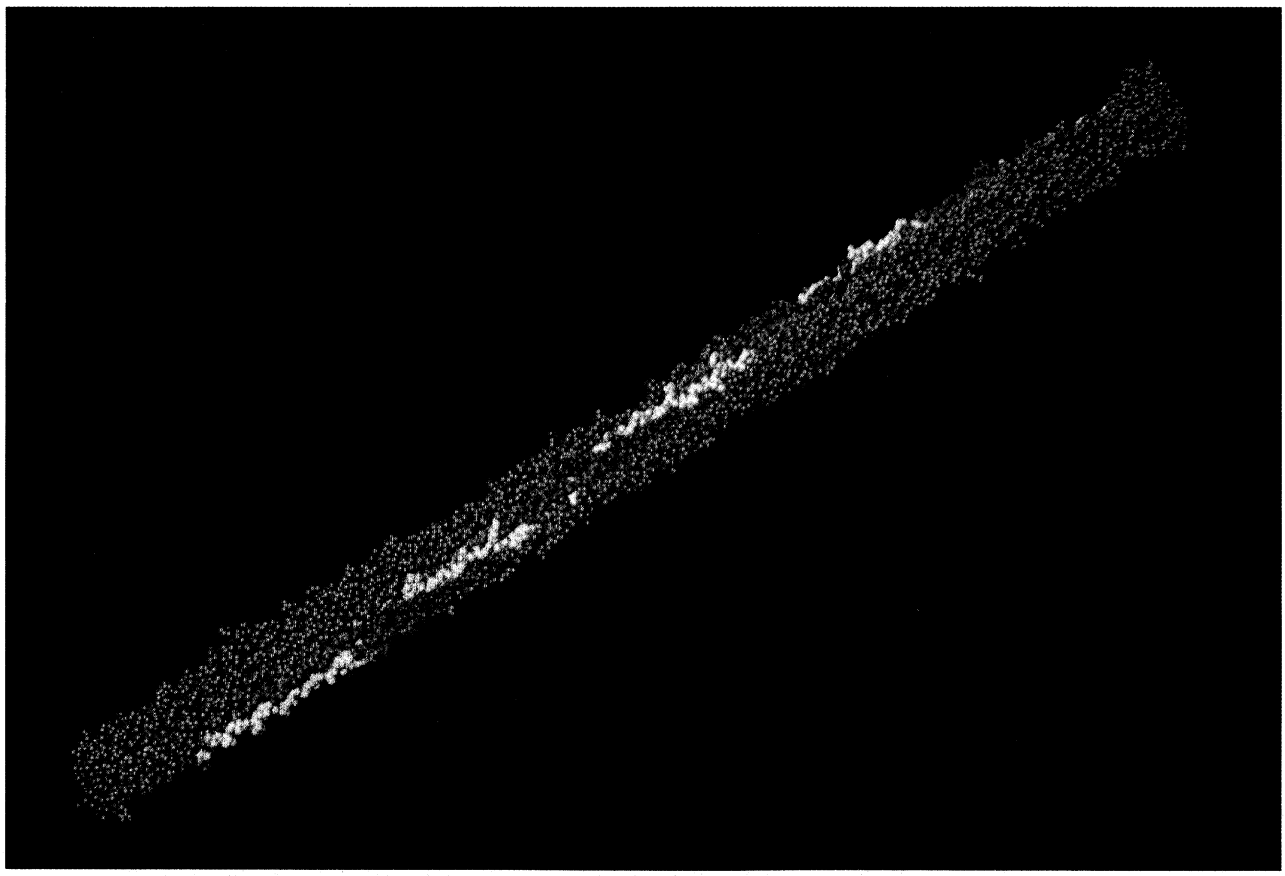

Fig. 3. A 400 Å segment of the microfibril model, colored to accentuate its rope-like organization. Four of the five triple helices are colored red. The remaining triple helix has its three constituent polypeptide strands colored green, blue and white. As the level of complexity increases from single polypeptide chain to triple helix to microfibril, the direction of the helices changes from left-handed to right-handed to left-handed.

to a supercomputer for further manipulation. Two microfibrils 5[3(Gly–Pro–Hyp)$_{69}$] were spliced together to produce a longer microfibril 5[3(Gly–Pro–Hyp)$_{105}$]. This microfibril model, consisting of 15 polypeptide chains each 315 residues long, was again energy minimized.

The structural scaffolding for the type I collagen microfibril model was complete with this addition, and the actual type I sequence could be substituted into the model. Care was taken in substituting the actual sequence into the model so as to minimize the disruption of the secondary structure. A different procedure from that used originally by Chen *et al.* (1991c) to substitute the actual sequence into the model was employed here. Chen *et al.* (1991c) changed the appropriate Pro and Hyp residues directly to their target residues, and then through visual inspection modified the positions of the protein side chains to correct those portions of the structure in which atoms overlapped. Because of the size of the current model, this strategy of correction by visual inspection would not have been efficient. Therefore we modified the side groups in a stepwise manner, such that at each modification the affected side groups were increased in size by at most one non-hydrogen atom. Thus the bulkier side chains were allowed to grow gradually into their sites within the protein in separate stages. This 'alchemy' (which should not be confused with the slow-growth method used in calculating free energy differences) is outlined in Figure 2, which shows that Pro or Hyp residues were converted to Ala in one step, and that four intermediate steps (Val, Leu, Met and Lys) were required to convert Pro or Hyp to Arg.

Results and discussion

The completed microfibril model 5[3(Gly–X–Y)$_{105}$] with the actual type I sequence is displayed in Figure 3. In this figure, the model has been colored to illustrate the rope-like qualities of the microfibril. At each level of organization, the direction of helical winding reverses. Each polypeptide chain (Gly–X–Y)$_{105}$ is in a left-handed polyproline type II helical conformation, with a pitch of ~3.3 residues. The triple helices 3(Gly–X–Y)$_{105}$ then switch to a right-handed helical conformation, with a pitch of ~28 residues. Then the microfibril 5[3(Gly–X–Y)$_{105}$] changes back to a left-handed superhelix, with a pitch of ~190 residues (550 Å).

Our current models do not include solvent, which plays an important role in moderating solute–solute interactions. The apparent stabilization energies we have calculated for the collagen triple helix relative to its three constituent polypeptide chains, and for the microfibril model relative to its five constituent triple helices, are thus not physically meaningful. The values obtained for these two stabilization energies were –15.4 and –7.0 k cal mol^{-1} residue^{-1}, respectively. These values are obviously overestimates of the true values because our calculations measure peptide–peptide energies in a vacuum, whereas the actual stabilization energies are a measure of peptide–peptide energies relative to peptide–water energies. In fact, the actual contribution to the stabilization energy from peptide–peptide hydrogen bonds may even be negligible, because the energetics of peptide–peptide and peptide–water hydrogen bonds are approximately the same. A similar statement may be made about van der Waals interactions.

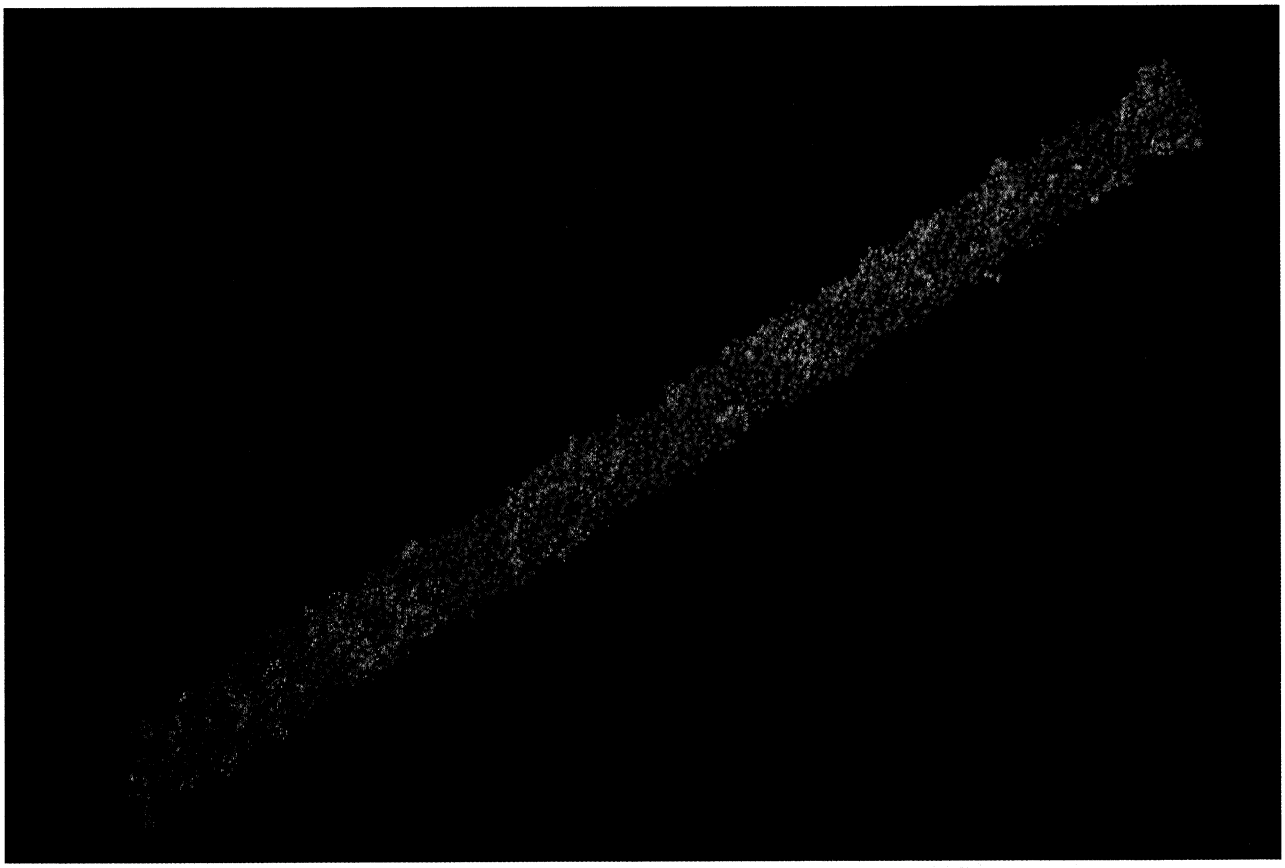

Fig. 4. The microfibril model is again shown, with the individual residues colored according to whether they are hydrophobic (green), polar (cyan), positively charged (purple) or negatively charged (red). The different types of residue are not distributed uniformly throughout the length of the molecule, but instead form regions of relatively high and low hydrophobicity, as well as regions of relatively high and low charge density.

In reality, the largest contribution to the stabilization of the microfibril relative to the triple helix, and of the triple helix relative to the individual peptide chain, is a result of the hydrophobic effect. The hydrophobic effect refers to phenomena in which the degree of organization of water is increased because of a reduction in the number of hydrogen bonding neighbors available to water molecules in the vicinity of nonpolar solutes. The hydrophobic effect is purely entropic; thus there is no term in the potential energy function that directly measures a system's hydrophobic energy, but there are empirical rules based on the solutes' solvent-accessible surface area (Nicholls *et al.*, 1991) for estimating a system's free energy.

The solvent-accessible surface areas of our models of an individual peptide chain, triple helix and microfibril are 119.0, 64.5 and 31.0 $Å^2$ $residue^{-1}$, respectively (calculated with a probe atom of radius 1.5 Å). The free energy difference between the triple-helical state and the individual peptide chain state is ~–2.7 kcal mol^{-1} $residue^{-1}$, using the microscopic surface tension of 0.05 kcal mol^{-1} $Å^{-2}$ (Nicholls *et al.*, 1991) to convert from solvent-accessible surface area to free energy. The free energy difference between the microfibril state and the triple-helical state is ~–1.7 kcal mol^{-1} $residue^{-1}$, and the free energy difference between the microfibril state and the individual peptide chain state is ~–4.4 kcal mol^{-1} $residue^{-1}$.

Figure 4 shows the model with the individual residues colored by type as follows: hydrophobic side chains (Ala, Ile, Leu, Met, Phe and Val) in green; neutral polar side chains (Asn, Gln, Ser and Thr) in cyan (this category also includes Gly, Pro and Hyp); positively charged side chains (Arg and Lys) in purple; and negatively charged side chains (Asp and Glu) in red. As can be seen from this figure, the different types of residues are not distributed uniformly throughout the length of the molecule, but instead form regions of relatively high and low hydrophobicity.

One goal of this research was to see if the computer model of the microfibril could reproduce the banding pattern observed in electron microscopy photographs of stained collagen (Figure 1). Each fibril is composed of several thousand microfibrils, presumably with every possible angular orientation with respect to the superhelical axis, and thus Figure 1 is representative of an orientationally averaged microfibril structure. To compare the computer model with this photograph, we attempted to create an orientationally averaged computer model. This was achieved by assigning a number to each atom in the following way: all atoms belonging to charged side groups were initially assigned a value of 1, and the rest of the atoms were initially assigned a value of 0. A filtering operation, in which all atoms within the same cross-sectional slice were given the same numerical index (where the index is constrained to take on values between 0 and 1), was then carried out. The resulting indices were interpreted as shades of gray so as to provide a direct comparison with Figure 1. As can be seen when Figure 1 is compared with Figure 5, the computer model displays a banding pattern similar to that seen in the electron microscopy image of chicken type I collagen (Mould *et al.*, 1990). The banding pattern displayed in Figure 5 is also shown

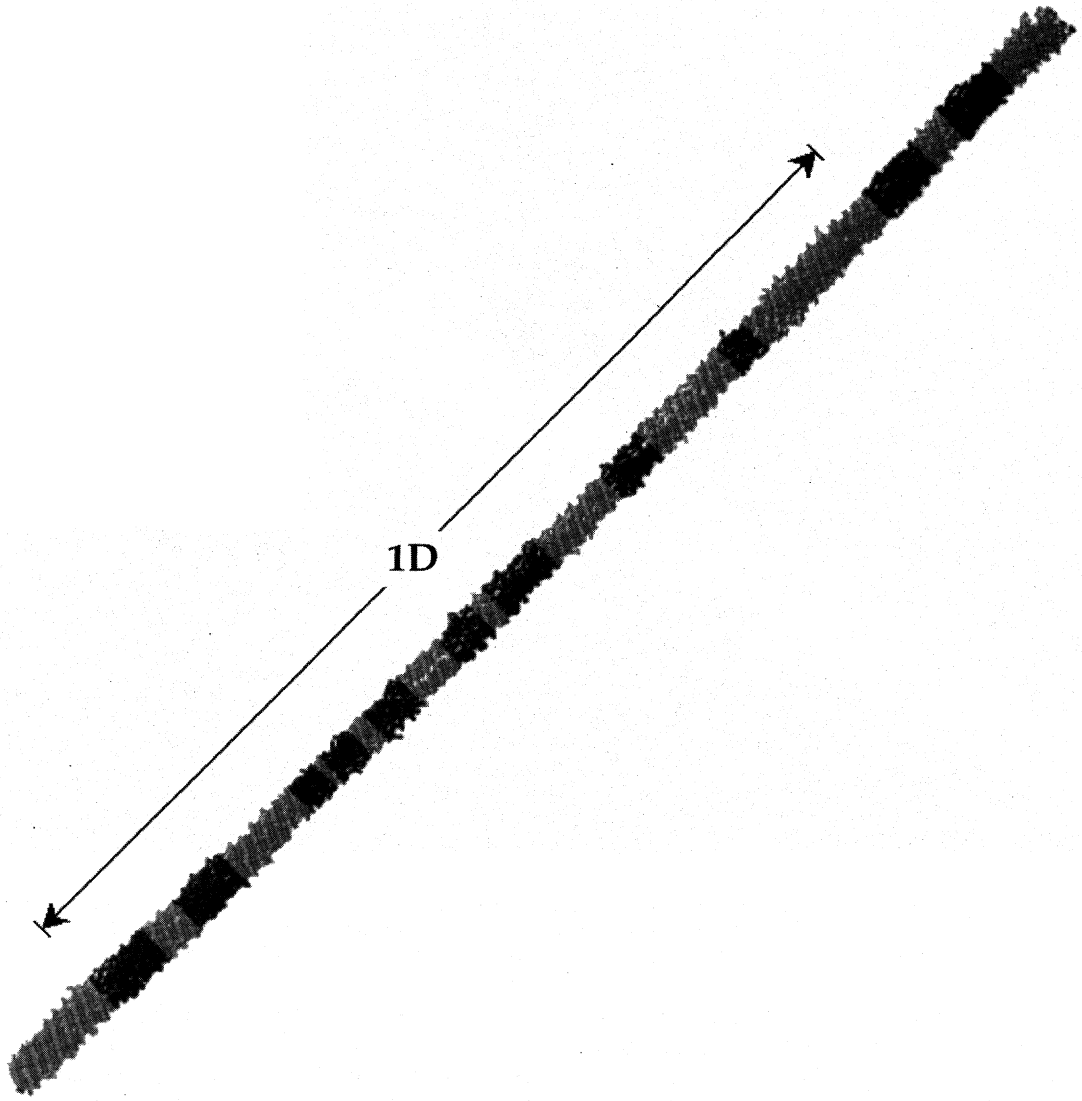

Fig. 5. Illustration of the computer model of the microfibril 5[3(Gly–X–Y)$_{105}$], rendered to simulate the banding observed in actual stained collagen fibrils. The ionized side groups have been given a different color from the rest of the protein. A color filtering operation which assigns the same color to all atoms within the same cross-sectional slice was carried out to compare more easily the model to the electron microscopy image of stained collagen fibrils (see Figure 1d).

as an inset in Figure 1d on the same scale as the electron microscopy image. The agreement between the experimental banding pattern and that obtained from our computer model does not specifically validate the Smith microfibril model because the banding pattern depends only on the fact that the triple-helical molecules are staggered with respect to one another, and not on the grouping of these molecules into larger structures.

Future studies using the microfibril model described here will examine the mechanisms involved in both intra-microfibril and inter-microfibril crosslinking. This model is expected to be useful in studies simulating the tanning of hides to make leather. Potential binding sites for a variety of chromium complexes (current tanning agents) and other types of crosslinking reagents will be explored with the model. All or parts of the model will be useful in evaluating the effects of point mutations on secondary, tertiary or quaternary structure. The use of smaller portions of the computer model in future studies will allow us to include water, which undoubtedly influences the structure and function of collagen, in the model system. When adding water to the system, it is conceivable that we may need to hydrate each of the five triple-helical molecules separately, and then rejoin them into a microfibrillar structure.

The model of the microfibril presented here will be made available in a standard ASCII format upon request to researchers wishing to evaluate it for other research projects. A minicomputer or workstation with at least 32 MB RAM and an appropriate molecular modeling software package will be necessary to use this model.

Acknowledgements

The authors thank Dr Stephen H.Feairheller for initiating this project. The mention of a brand or firm name does not constitute endorsement by the US Department of Agriculture over others of a similar nature not mentioned. This work was presented in part at the 1994 Meeting of the American Society for Biochemistry and Molecular Biology, Washington, DC.

References

Bailey,A.J. (1992) *J. Soc. Leather Tech. Chem.*, **76**, 111–127.

Brown,E.M., Chen,J.M. and Feairheller,S.H. (1993) *J. Am. Leather Chem. Assoc.*, **88**, 2–11.

Carver,J.P. and Blout,E.R. (1967) In Ramachandran,G.N. (ed.), *Treatise on Collagen.* Academic Press, London, UK, pp. 441–523.
Chen,J.M. and Sheldon,A. (1994) In Kumosinski,T.F. and Liebman,M.N. (eds), *Molecular Modeling: From Virtual Tools to Real Problems.* American Chemistry Society, Washington, DC, pp. 139–170.
Chen,J.M., Kung,C.E., Feairheller,S.H. and Brown,E.M. (1991a) *J. Protein Chem.*, **10**, 535–552.
Chen,J.M., Feairheller,S.H. and Brown,E.M. (1991b) *J. Am. Leather Chem. Assoc.*, **86**, 475–486.
Chen,J.M., Feairheller,S.H. and Brown,E.M. (1991c) *J. Am. Leather Chem. Assoc.*, **86**, 487–497.
Fietzek,P.P. and Kuhn,K. (1976) *Int. Rev. Connective Tissue Res.*, **7**, 1–60.
Fraser,R.D.B., Miller,A. and Parry,D.A.D. (1974) *J. Mol. Biol.*, **83**, 281–283.
Fraser,R.D.B., MacRae,T.P. and Suzuki,E. (1979) *J. Mol. Biol.*, **129**, 463–481.
Fraser,R.D.B., MacRae,T.P., Miller,A. and Suzuki,E. (1983) *J. Mol. Biol.*, **167**, 497–521.
Hodge,A.J. and Petruska,A.J. (1963) In Ramachandran,G.N. (ed.), *Aspects of Protein Structure.* Academic Press, London, UK, pp. 289–300.
Hofmann,H., Fietzek,P.P. and Kuhn,K. (1978) *J. Mol. Biol.*, **125**, 137–165.
Miller,M.H. and Scheraga,H.A. (1976) *J. Polymer Sci., Polym. Symp.*, **54**, 171–200.
Mould,A.P., Hulmes,D.J.S., Holmes,D.F., Cummings,C., Sear,C.H.J. and Chapman,J.A. (1990) *J. Mol. Biol.*, **211**, 581–594.
Nicholls,A., Sharp,K.A. and Honig,B. (1991) *Proteins: Struct. Funct. Genet.*, **11**, 281–296.
Okuyama,K., Tanaka,N., Ashida,T., Kakudo,M., Sakakibara,S. and Kishida,Y. (1972) *J. Mol. Biol.*, **72**, 571–576.
Okuyama,K., Tanaka,N., Ashida,T. and Kakudo,M. (1976) *Bull. Chem. Soc. Japan*, **49**, 1805–1810.
Okuyama,K., Go,N. and Takayanagi,M. (1978) *Chem. Lett.*, 509–512.
Piez,K.A. and Trus,B.L. (1977) *J. Mol. Biol.*, **110**, 701–704.
Piez,K.A. and Trus,B.L. (1978) *J. Mol. Biol.*, **122**, 419–432.
Ramachandran,G.N. (1967) In Ramachandran,G.N. (ed.), *Treatise on Collagen.* Academic Press, London, UK, pp. 103–183.
Ramachandran,G.N. (1968) *J. Am. Leather Chem. Assoc.*, **63**, 160–191.
Ramachandran,G.N. and Kartha,G. (1954) *Nature*, **174**, 269–270.
Rich,A. and Crick,F.H.C. (1961) *J. Mol. Biol.*, **3**, 483–506.
Rose,P.I. (1992) In Pearson,A.M. and Dutson,T.R. (eds), *Inedible Meat By-products, Advances in Meat Research.* Elsevier, New York, pp. 217–263.
Sakakibara,S., Kishida,Y., Kikuchi,Y., Sakai,R. and Kakiuchi,K. (1968) *Bull. Chem. Soc. Japan*, **41**, 1273.
Scholnick,F., Liao,L.L., Brown,E.M. and Feairheller,S.H. (1992) *J. Am. Leather Chem. Assoc.*, **87**, 333–338.
Smith,J.W. (1968) *Nature*, **219**, 157–158.
Traub,W. (1978) *FEBS Lett.*, **92**, 114–120.
Van der Rest,M. and Garrone,R. (1991) *FASEB J.*, **5**, 2814–2823.
Veis,A. and Yuan,L. (1975) *Biopolymers*, **14**, 895–900.
Weiner,P.K. and Kollman,P.A. (1981) *J. Comp. Chem.*, **2**, 287–303.
Weiner,S.J., Kollman,P.A., Case,D.A., Singh,U.C., Ghio,C., Alagona,G., Profeta,S. and Weiner,P.A. (1984) *J. Am. Chem. Soc.*, **106**, 765–784.